ADVANCED TRAFFIC MODELLING FOR SMART CITIES USING POWERFUL DATA PROCESSING PLATFORM

AUTHORS: RADIM CMAR (SYGIC), MARTIN ŠURKOVSKÝ, KATEŘINA SLANINOVÁ, JAN MARTINOVIČ (IT4INNOVATIONS)

Concept and approach

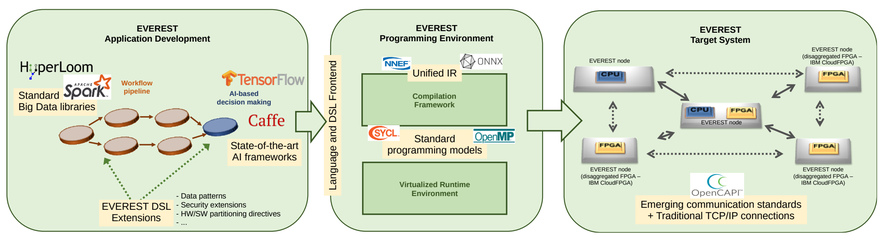
EVEREST focuses on **High Performance Big Data Analytics (HPDA) applications.**

- Future Big Data systems will be data-driven.
- Complex heterogeneous and reconfigurable architectures are difficult to program.

The EVEREST project aims at developing a holistic approach for co-designing computation and communication in a heterogeneous, distributed, scalable, and secure system for HPDA.

Main features:

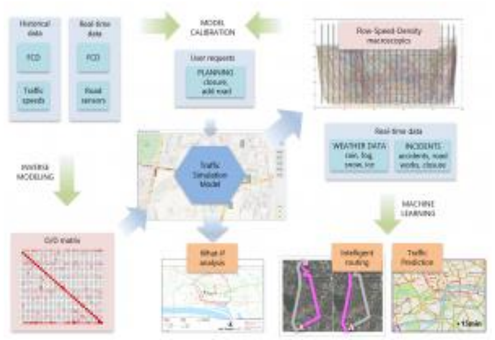
- **data-driven design approach;**
- combination of **compiler transformations, high-level synthesis, and memory management;**
- efficient monitoring of the execution with a **virtualisation-based environment.**



EVEREST proposes a **design environment** that combines state-of-the-art, stable programming models, and emerging communication standards with novel and **dedicated domain-specific extensions.** The EVEREST approach will be validated on three industrial use cases, one is related to **advanced traffic modelling for smart cities.**

Application use case: Advanced traffic modelling for smart cities

Sygic provides the **mobility platform** for supporting cities with advanced traffic modelling. The platform absorbs **big data** and deploys **traffic services.**



Data sources

- **Historical** and **real-time Floating car data** (FCD) from navigation devices defining GPS position, timestamp, speed and bearing (typically with 5 second period);
- **origin-destination matrix** (ODM) defining mobility of daily commuters across the city grid;
- **road network graph** including **road restrictions;**
- **historical weather data** (temperature, precipitation).

Traffic services

- **What-if traffic analysis** for given hypothetical scenario, such as road closure;
- **intelligent routing** for large amount of vehicles **towards a global optimum;**
- **traffic prediction** for major road elements of cities.

The platform makes use of **IT4Innovations** traffic simulator, which boosts the FCD data into an extended set to compute **3D traffic model**. Platform applies **AI** techniques to learn the traffic patterns resulting into traffic prediction service.

Data processing in traffic modelling

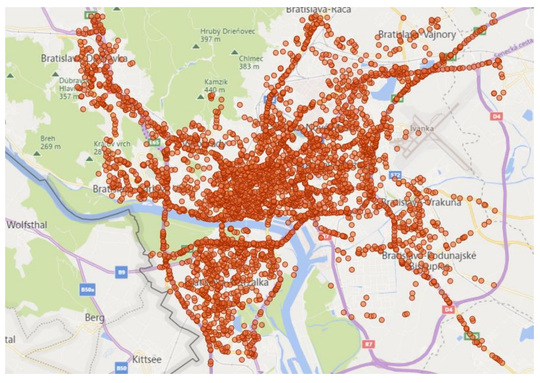
Big data

Floating car data (FCD):

- 500 million points in Europe daily;
- Vienna region – 2 million points daily;
- infers speed profiles for each road.

Origin-destination matrix (ODM):

- 500 x 500 meter cell grid;
- mobility flow count on 15min interval throughout weekdays between each grid cell pair.

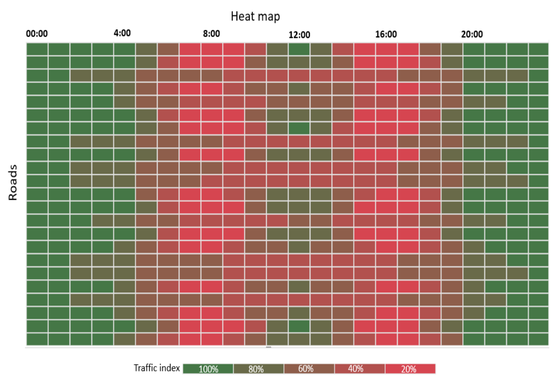


AI algorithms on traffic data sets

Big data combined with AI algorithms allow building city traffic prediction model

- **Neural Networks** (deep, convolutional, recurrent) are used with time sequences to learn and infer 1 hour traffic prediction for each major road in city based on actual state and short history.
- **Gaussian Mixture Model** is used for traffic prediction using Bayesian inference where some input data are missing or are incomplete.
- **Hidden Markov Model** is the pre-processing step to boost sparse FCD data points into reach road speed annotations.

For facilitating data application designs EVEREST offers many AI library components **optimizable** for use with C++ and Python **compilers, high-level synthesis** for reconfigurable **FPGAs**, and with **specialised memory managers.**



EVEREST data management methodology

Everest deals with data management with several aspects:

- **Performance** – data transfers among processing elements are mapped on fast technologies.
- **Scalability** – architecture can expand accordingly with data size.
- **Security** – data are protected for unauthorized access and secured for tempering.

Data boost technology

Traffic simulator by IT4Innovations

- Using ODM and FCD data the city traffic view can be created. This generates missing speed points across the whole network.

Map matching with Dijkstra Routing

- Using map matching algorithm (Hidden Markov Model) speeds on all road segments on a trajectory can be inferred from sparse sequence of FCD points.

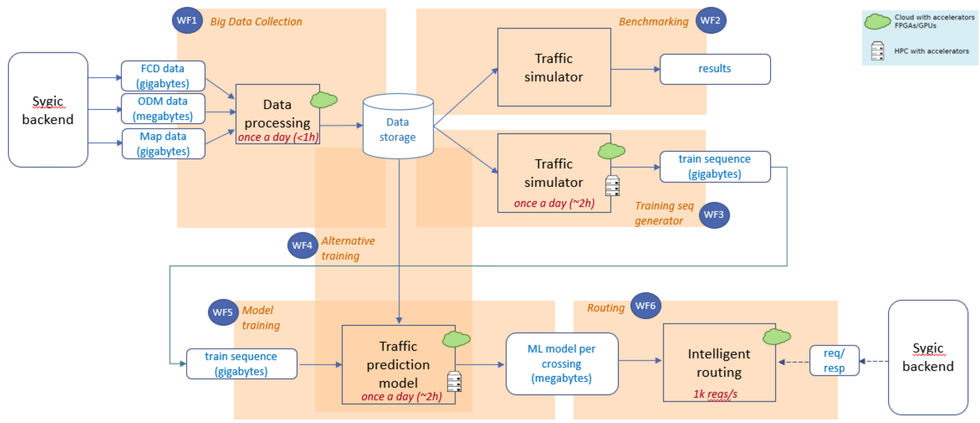
Data storage implementation

- GBs/TBs data sequences for model training are stored on HPC premises.
- **FPGA** accelerators used for inference calculation use their GB local storage to keep trained model state.

Data processing with EVEREST environment

EVEREST is targeting FPGA-based distributed architectures to accelerate extreme-scale big data applications.

- System is split into **workflows,** where the workflow is a complex architecture within a single compute centre. Batch-oriented workflows are executed on HPC infrastructure, streaming-oriented workflows are executed on a cloud infrastructure.
- Critical **kernels** are mapped to **FPGA based architectures.** Examples: convolution neural network (CNN), gaussian mixture model (GMM), Hidden Markov Model (HMM). FPGA systems are generated with high-level synthesis optimised for data movements.
- EVEREST provides **data-driven compilation framework** with orchestrated data programming flow using **HyperLoom** tooling, kernel accelerations with high-level semantics on domain-specific languages, and ML programming support through standard libraries, e.g. **Tensorflow.**



EVEREST
DESIGN ENVIRONMENT FOR EXTREME-SCALE BIG DATA ANALYTICS ON HETEROGENEOUS PLATFORMS

WWW.EVEREST-H2020.EU
EVEREST-INFO@A.ALARI.CH
WWW.LINKEDIN.COM/COMPANY/EVEREST-H2020
WWW.FACEBOOK.COM/EVERESTH2020

PROJECT COORDINATOR: CHRISTOPH HAGLEITNER (IBM ZURICH) | **SCIENTIFIC COORDINATOR:** CHRISTIAN PILATO (POLITECNICO DI MILANO)

GRANT NUMBER: 957269 | TOPIC: ICT-51-2020 – BIG DATA TECHNOLOGIES AND EXTREME-SCALE ANALYTICS | FUNDING SCHEME: RIA